EUROPEAN PATENT OFFICE

Patent Abstracts of Japan

PUBLICATION NUMBER : 08320768
PUBLICATION DATE : 03-12-96

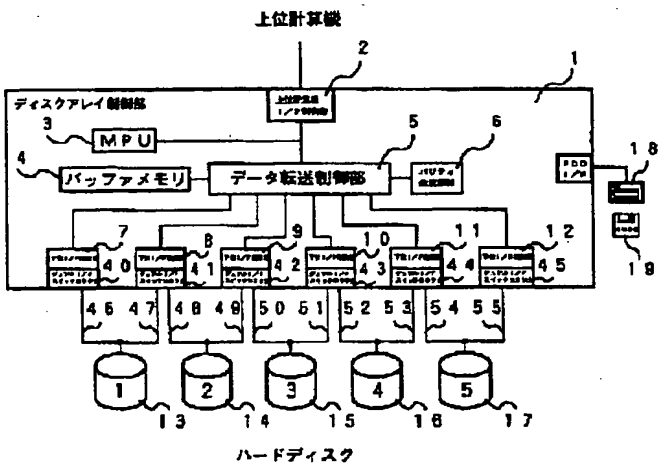
APPLICATION DATE : 26-05-95
APPLICATION NUMBER : 07127761

APPLICANT : HITACHI LTD;

INVENTOR : IWASAKI HIDEHIKO;

INT.CL. : G06F 3/06 G06F 3/06

TITLE : DISK ARRAY DEVICE



ABSTRACT : PURPOSE: To provide the disk array device which is placed in continuous operation if a fault occurs to one slave I/F control part or the cable connecting the slave I/F control part and a hard disk.

CONSTITUTION: Between slave I/F control parts 7-12 of a disk array control part 1 and hard disks 13-17, dual I/F connectors 40-45 are arranged which switch the main path using cables 47, 49, 51, 53, and 55 used when the disk array device is in normal operation as communication paths to a standby path using cables 46, 48, 50, 52, and 54 used if a fault occurs to one slave I/F control part or the cable connecting the slave I/F control part and a hard disk as communication paths; and each hard disk has two communication means for the slave I/F control parts 7-12.

(19)日本国特許庁（ＪＰ）　(12) 公開特許公報（A）　(11)特許出願公開番号

特開平8－320768

(43)公開日　平成８年(1996)12月３日

(51)Int.Cl.[6]	識別記号	庁内整理番号	ＦＩ		技術表示箇所
Ｇ０６Ｆ　3/06	５４０		Ｇ０６Ｆ　3/06	５４０	
	３０４			３０４Ｈ	

審査請求　未請求　請求項の数10　ＯＬ　（全 24 頁）

(21)出願番号　特願平7－127761

(22)出願日　平成７年(1995)５月26日

(71)出願人　000005108
株式会社日立製作所
東京都千代田区神田駿河台四丁目６番地

(72)発明者　松本　純
神奈川県川崎市麻生区王禅寺1099番地株式会社日立製作所システム開発研究所内

(72)発明者　大枝　高
神奈川県川崎市麻生区王禅寺1099番地株式会社日立製作所システム開発研究所内

(72)発明者　市川　正敏
神奈川県川崎市麻生区王禅寺1099番地株式会社日立製作所システム開発研究所内

(74)代理人　弁理士　小川　勝男

最終頁に続く

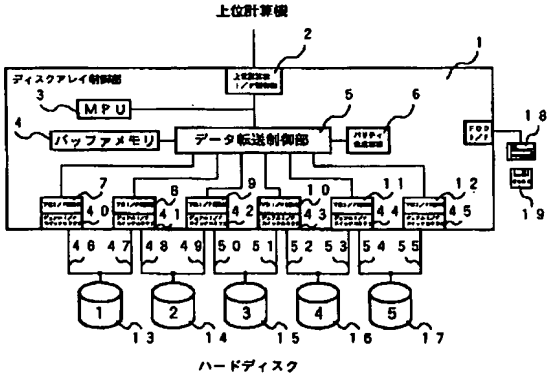
(54)【発明の名称】　ディスクアレイ装置

(57)【要約】

【目的】１台の下位Ｉ／Ｆ制御部または下位Ｉ／Ｆ制御部とハードディスクを接続するケーブルの故障時に、連続運用可能なディスクアレイ装置を提供する。

【構成】ディスクアレイ制御部１の下位Ｉ／Ｆ制御部７～１２とハードディスク１３～１７の間に、ディスクアレイ装置の正常運用時に用いケーブル４７、４９、５１、５３、５５を通信経路とする主経路から、１台の下位Ｉ／Ｆ制御部または下位Ｉ／Ｆ制御部とハードディスクを接続するケーブルの故障時に用いケーブル４６、４８、５０、５２、５４を通信経路とする予備経路に切替えるデュアルＩ／Ｆコネクタ４０～４５を備え、各々のハードディスクが下位Ｉ／Ｆ制御部７～１２に対して２本の通信手段を持つ。

図 1



1

【特許請求の範囲】

【請求項１】上位に上位計算機、下位に複数のハードディスクを接続し、内部に上位計算機との間でコマンド及びデータの授受を行う上位計算機Ｉ／Ｆ制御部、複数のハードディスクとの間でコマンド及びデータの授受を行う下位Ｉ／Ｆ制御部、コマンドのアドレス変換、データの集合分散制御を行うデータ転送制御部、データからパリティの生成を行うパリティ生成回路、上位計算機と複数のハードディスクの間で転送するデータを一時的に格納保持するバッファメモリ、さらに以上の各構成要素を統括制御するＭＰＵを内部に持つディスクアレイ制御部において、Ｎ台の下位Ｉ／Ｆ制御部＋１台の予備用の下位Ｉ／Ｆ制御部を持つことで、１台の下位Ｉ／Ｆ制御部の故障発生時に連続運用が可能なことを特徴とするディスクアレイ装置。

【請求項２】請求項１において、１台の下位Ｉ／Ｆ制御部の故障発生時に、下位Ｉ／Ｆ制御部とハードディスク間の通信経路を正常運用時に使用する主経路から、下位Ｉ／Ｆ制御部及びディスクアレイ制御部とハードディスクの間を接続するケーブルの故障発生時に使用する予備経路に切替えることで、連続運用が可能なディスクアレイ装置。

【請求項３】請求項２において、各下位Ｉ／Ｆ制御部と各ハードディスク間に正常運用時に使用する主経路と、下位Ｉ／Ｆ制御部及び下位Ｉ／Ｆ制御部とハードディスクを接続するケーブルの故障発生時に使用する予備経路を切替る通信経路切替スイッチを付加したディスクアレイ装置。

【請求項４】請求項３において、１台の下位Ｉ／Ｆ制御部の故障発生時にディスクアレイ制御部を構成する基盤全体ではなく、下位Ｉ／Ｆ制御部単位で故障した下位Ｉ／Ｆ制御部の交換を可能とするため、各々の下位Ｉ／Ｆ制御部をソケットにより抜き差し可能な構造を持つディスクアレイ装置。

【請求項５】請求項４において、外部機器専用の下位Ｉ／Ｆ制御部を持たずにＭＴ、ＭＯ等のバックアップ機器を接続可能なディスクアレイ装置。

【請求項６】請求項５において、１台の下位Ｉ／Ｆ制御部の故障発生時にも残りの下位Ｉ／Ｆ制御部を利用してＭＴ、ＭＯ等のバックアップ機器にデータバックアップ可能なディスクアレイ装置。

【請求項７】請求項４に記載の通信経路切替スイッチを搭載し、従来のディスクアレイ装置の下位Ｉ／Ｆ制御部とハードディスク間のケーブルの途中に接続されるドータボード。

【請求項８】請求項８に記載のドータボードを、従来のディスクアレイ装置の下位Ｉ／Ｆ制御部とハードディスク間に付加したディスクアレイ装置。

【請求項９】上位に上位計算機、下位に複数のハードディスクを接続し、内部に上位計算機との間でコマンド及びデータの授受を行う上位計算機Ｉ／Ｆ制御部、複数のハードディスクとの間でコマンド及びデータの授受を行うシリアルＩ／Ｆ制御部、コマンドのアドレス変換、データの集合分散制御を行うデータ転送制御部、データからパリティの生成を行うパリティ生成回路、上位計算機と複数のハードディスクの間で転送するデータを一時的に格納保持するバッファメモリ、さらに以上の各構成要素を統括制御するＭＰＵを内部に持つディスクアレイ制御部において、ディスクアレイ制御部のシリアルＩ／Ｆ制御部をデュアル化したことを特徴とするディスクアレイ装置。

【請求項１０】請求項９において、２台のシリアルＩ／Ｆ制御部と各ハードディスク間に正常運用時に使用する主経路と、下位Ｉ／Ｆ制御部及び下位Ｉ／Ｆ制御部とハードディスクを接続するケーブルの故障発生時に使用する予備経路を切替る通信経路切替スイッチを付加したディスクアレイ装置。

【発明の詳細な説明】

【０００１】

【産業上の利用分野】本発明は、ハードディスク、光ディスク、半導体ディスクなどのディスク装置を複数使用するディスク装置システムにおけるディスクアレイ装置に関する。

【０００２】

【従来の技術】ファイルシステムの大容量化と高いトランザクション性能が要求される中、高いデータ処理性能と信頼性を併せ持つ一方式として、米国特許ＵＳ４８７０６４３号に開示されるパラレルドライブアレイ・ストレージシステムがある。本方式は従来のディスクアレイ装置の基本構成を示すものである。ここで従来のディスクアレイ装置について図１６を用いて説明する。

【０００３】ディスクアレイ装置は主に上位計算機から受けたリード／ライトコマンドにより、上位計算機とハードディスクの間でデータの分散集合制御を行うディスクアレイ制御部１と上位計算機が転送したデータを格納するハードディスク１３～１７で構成される。またデータ保持の信頼性を高める意味で、ハードディスク１３～１７のデータをバックアップするＭＴやＭＯ等のバックアップ機器をオプションとして接続することが多い。ディスクアレイ制御部１は上位計算機との間でコマンド及びデータの授受を行う上位計算機Ｉ／Ｆ制御部２、各ハードディスク１３～１７及びバックアップ機器２１との間でコマンド及びデータの授受を行う下位Ｉ／Ｆ制御部７～１２、上位計算機Ｉ／Ｆ制御部２と下位Ｉ／Ｆ制御部７～１１の間でデータを転送制御するデータ転送制御部５、ハードディスク１３～１７に新しいデータをライトする際にデータに付加するパリティを生成するパリティ生成回路６、また上位計算機とハードディスク１３～１７間のデータ転送の際に、一時的にデータを保持するバッファメモリ４、さらにディスクアレイ制御部１全体

を統括制御するＭＰＵ３等で構成される。本ディスクアレイ装置は各ハードディスク１３～１７それぞれに下位Ｉ／Ｆ制御部７～１１を備えるため、複数ハードディスク１３～１７のパラレルなリード／ライト動作が可能となり、単体のハードディスクに比べて特にアクセス性能に優れているという特徴を持つ。また、データをハードディスクにライトする際に、パリティ生成回路６が図１７に示すようにデータの排他的論理和からパリティを生成し、データと共にハードディスクのある場所にライトするため、例えば、３番目のハードディスク１５に故障が発生しても、残りのハードディスク１３、１４、１６、１７のデータＤ１、Ｄ２、Ｄ４、Ｄ５及びパリティＰから喪失データＤ３を再生することができる。

【０００４】ハードウエア構成を持つディスクアレイ装置は、下位Ｉ／Ｆ制御部７～１２に特にＳＣＳＩ<Small Computer System Interface>を用いたもので、ＳＣＳＩはコンピュータ周辺機器の一般的なＩ／Ｆ制御部として広く普及している。しかし、近年ＭＰＵ、メモリ及びハードディスク等のハードウエアの性能の上昇や小型化が著しいため、ＳＣＳＩの持つデータ転送速度や転送距離による制約、大型のＩ／Ｆコネクタが問題となっており、これに変わるものとしてシリアルＩ／Ｆ制御部が注目されている。図１８はハードディスク２４～２８とディスクアレイ制御部２９の間にシリアルＩ／Ｆ制御部２２を用いたディスクアレイ装置を示した図である。ディスクアレイ制御部２９を構成する要素はシリアルＩ／Ｆ制御部２２を除いてＳＣＳＩＩ／Ｆ制御部を用いたディスクアレイ装置と同様であるが、シリアルＩ／Ｆ制御部２２の高い転送能力を引き出すために、専用のシリアルＩ／Ｆ制御用ＭＰＵ２３を持っている。またハードディスク２４～２８は各々のデータ入出力口に、ディスクアレイ制御部２９同様にシリアルＩ／Ｆ制御部３０～３４を備えている。図１９はディスクアレイ制御部２９側のシリアルＩ／Ｆ制御部２２（特にＳＳＡ<Serial Storage Architecture>）とハードディスク２４～２８側のシリアルＩ／Ｆ制御部３０～３４の詳細構成を示している。シリアルＩ／Ｆ制御部２２、３０～３４はデータの送信及び受信用ポートを２ポートずつ備えており、各シリアルＩ／Ｆ制御部２２、３０～３４がループ状に連なることで、隣り合うシリアルＩ／Ｆ制御部が双方向の通信経路を持つことができる。このことから例えば図２０に示すように、上位計算機からハードディスク２４と２７にデータをライトすると同時に、ハードディスク２５と２８からデータをリードし、ハードディスク２６からハードディスク２７にデータをバックアップするといったことが可能になる。またハードディスク２５に故障が発生した場合には例えば図２１に示すように、、ハードディスク２４にはディスクアレイ制御部側との送信及び受信用ポートが、またハードディスク２６にはハードディスク２７との送信及び受信用ポートが残っている。

したがって、上位計算機からは故障したハードディスクを除いて、残りの全てのハードディスクに対してリード／ライトアクセスが行えるため、ディスクアレイ装置の連続運用が可能である。

【０００５】

【発明が解決しようとする課題】従来の方法では、以下の課題がある。

【０００６】すなわち、まず下位Ｉ／Ｆ制御部にＳＣＳＩを用いるディスクアレイ装置について述べる。ディスクアレイ制御部１が有する下位Ｉ／Ｆ制御部に１台以上の故障が発生した場合、その下位Ｉ／Ｆ制御部と接続しているハードディスクに対する上位計算機からのリード／ライトアクセスが完全に遮断されることが考えられる。また下位Ｉ／Ｆ制御部とハードディスクを相互に接続しているケーブルに、断線等の故障が発生した場合にもハードディスクに対するアクセスが停止すると考えられる。したがってこのような条件下では、故障の発生したＩ／Ｆチャネルを除いた構成でのディスクアレイ装置の連続運用は可能であるが、１台のハードディスクが使えないことによって、正常運用時に比較してディスクアレイ装置のリード／ライトのアクセス性能が低下すると予想される。

【０００７】一方、下位Ｉ／Ｆ制御部にシリアルＩ／Ｆ制御部を用いるディスクアレイ装置について述べると、ディスクアレイ制御部及び各ハードディスクの持つ各々のシリアルＩ／Ｆ制御部がループ状に連なるため、ディスクアレイ制御部からハードディスクへのアクセス経路は、送信用及び受信用を１組にして２系統を保持している。これによって、いずれのハードディスクに故障が発生しても、残りの正常なハードディスクに対してアクセスを継続可能であるため、装置として高い信頼性がある。しかし、ディスクアレイ制御部側のシリアルＩ／Ｆ制御部に故障が発生した場合には、上位計算機とハードディスク間の通信経路を失うことから、ハードディスクのリード／ライトアクセスが事実上不可能となる。またディスクアレイ制御部側のシリアルＩ／Ｆ制御部と、ハードディスク側のシリアルＩ／Ｆ制御部間のケーブルに断線等の故障が一つ発生すると、２系統ある通信経路の内一つの送信用もしくは受信用の通信経路を失うことから、正常運用時に比較してディスクアレイ装置のリード／ライトのアクセス性能が低下すると予想される。

【０００８】

【課題を解決するための手段】上記の目的を達成するため、ディスクアレイ制御部に予備の下位Ｉ／Ｆ制御部を搭載し、各下位Ｉ／Ｆ制御部が異なる２台ハードディスクと通信経路を持つようにデュアル化した通信手段（主にケーブルを意味する）、および各下位Ｉ／Ｆ制御部と通信経路の間に正常時に使用する主経路と故障時に使用する予備経路を選択可能な通信経路切替手段を設ける。

【０００９】

5

【作用】上記の目的を達成する方法を用いると、下位Ｉ／Ｆ制御部及び通信経路に故障が発生していない場合、通信経路切替手段の主経路を有効、予備経路を無効に設定し、主経路と接続した通信経路の末端に接続してあるハードディスクに対してリード／ライトアクセスを行う。またある１台の下位Ｉ／Ｆ制御部が故障で使用不能になった場合、または運用中に通信経路が断線した場合には、故障した下位Ｉ／Ｆ制御部または断線した通信経路を持つ下位Ｉ／Ｆ制御部に対して右または左に隣り合う、全ての下位Ｉ／Ｆ制御部の通信経路切替手段の主経路を無効、予備経路を有効に設定し、予備経路と接続した通信経路の末端に接続してあるハードディスクに対してリード／ライトアクセスを継続して行う。ディスクアレイ装置の運用はこの状態で継続可能である。そして運用の終了後、一度ディスクアレイ装置の電源を落とし、故障した下位Ｉ／Ｆ制御部を新しい下位Ｉ／Ｆ制御部に交換する。この時、交換した下位Ｉ／Ｆ制御部に対して左に隣り合う、全ての下位Ｉ／Ｆ制御部の通信経路切替手段の主経路を再度有効、予備経路を無効に設定し、主経路と接続した通信経路の末端に接続してあるハードディスクに対してリード／ライトアクセスを行う。尚通信経路切替手段を別のドータボードに搭載し、ディスクアレイ制御部を構成する基盤上の下位Ｉ／Ｆ制御部とハードディスクの間にそのドータボードを付加することで、従来のディスクアレイ装置に対しても本方式の適用が図れる。

【００１０】

【実施例】本発明の第一実施例を図１ないし図８を用いて説明する。

【００１１】図１は本発明のディスクアレイ装置全体を表した図である。ディスクアレイ装置の中枢となるディスクアレイ制御部１は、上位計算機との間でコマンド及びデータの授受を行う上位計算機Ｉ／Ｆ制御部２、複数のハードディスク１３～１７との間でコマンド及びデータの授受を行う下位Ｉ／Ｆ制御部７～１２、各々の下位Ｉ／Ｆ制御部７～１２と相互に接続し、下位Ｉ／Ｆ制御部とハードディスク間の通信経路（以下ケーブルと記す）４６～５５の有効／無効の切替えを行うデュアルＩ／Ｆスイッチコネクタ４０～４５、コマンドのアドレス変換及びデータの集合分散制御を行うデータ転送制御部５、上位計算機とハードディスク１３～１７間で転送するデータを一時的に格納保持するバッファメモリ４、ハードディスク１３～１７にデータをライトする際にデータに付加するパリティを生成するパリティ生成回路６、さらに以上の各構成要素を統括制御するＭＰＵ３から構成される。尚デュアルＩ／Ｆスイッチコネクタ４０はハードディスク１３に、デュアルＩ／Ｆスイッチコネクタ４５はハードディスク１７に対して１本の通信経路を備え、残りのデュアルＩ／Ｆスイッチコネクタ４１～４４はハードディスク１３～１７に対して各々２本の通信経路を備えている。また各ハードディスクは図８に示すようなケーブルを用いて、隣り合う２台のデュアルＩ／Ｆスイッチコネクタの各々１つの通信ポートと接続している。さらにディスクアレイ装置運用開始時にディスクアレイ制御部１上の各ハードウエアをチェックするＣＵＤＧ<Controller Unit Diagnostic>プログラムをフロッピディスク１９で供給するため、ディスクアレイ制御部１にフロッピディスクドライブ（以下ＦＤＤと記す）１８を接続している。

6

【００１２】図２～図４のフローチャートに基づき本発明のディスクアレイ装置の動作について説明する。

【００１３】ディスクアレイ装置の電源を入れると（図２のフローチャート１００）、ＦＤＤ１８に入っているフロッピディスク１９からロードされたＣＵＤＧプログラムが起動し、ディスクアレイ制御部１の各構成要素である上位計算機Ｉ／Ｆ制御部２、ＭＰＵ３、バッファメモリ４、データ転送制御部５、パリティ生成回路６、各下位Ｉ／Ｆ制御部７～１２及び各ハードディスク１３～１７のハードチェックを行う（同１０１）。この時下位Ｉ／Ｆ制御部７～１２のいずれかの障害の発生、または下位Ｉ／Ｆ制御部７～１２とハードディスク１３～１７をつなぐケーブルの断線によって、ハードディスク１３～１７のハードチェックが正常に行われない場合は（同１０２）、ＭＰＵ２がエラー情報を上位計算機に報告する（同１０３）。運用者はディスクアレイ装置の電源をいったん落とし、故障の認められた下位Ｉ／Ｆ制御部を交換するか、または断線していると思われるケーブルを交換する（同１０５）。一方ディスクアレイ装置の電源投入時のハードチェックにおいて、いかなるハードエラーも認められなかった場合は（同１０２）、図５の上側に示す下位通信経路の接続構成でディスクアレイ装置の運用を開始する。上位計算機は運用者の要求に応じてリード／ライトコマンドを送出し（同１０６）、ディスクアレイ制御部１のＭＰＵ２がコマンドを解釈し、データ転送制御部５がリード時には各ハードディスク１３～１７に分散したデータの読み出し集合処理、ライト時には各ハードディスク１３～１７へデータの分散書き込み処理を実行する（同１０７）。ディスクアレイ装置のリード／ライトアクセス実行中にいずれの下位Ｉ／Ｆ制御部、ケーブル及びハードディスクの故障がなく（同１０８）、実行中のリード／ライトアクセスがまだ終了していない場合は（同１０９）、引続きリード／ライトアクセスを実行する（同１０７）。一方リード／ライトアクセスが終了している場合（同１０９）、ディスクアレイ制御部１が上位計算機から次のリード／ライトアクセス要求を受け付ける。次のリード／ライトアクセス要求があるならば（同１１０）上位計算機が次のリード／ライトコマンドをディスクアレイ制御部１に対して送出し（同１０６）、ないならば（同１１０）ディスクアレイ装置の運用を終了する（同１１１）。

7

【００１４】ここで例えば図１の下位Ｉ／Ｆ制御部９、ケーブル４９またはハードディスク１４のいずれかに故障が発生したと仮定する（同１０８）。リード時には（同１１２）データ転送制御部５が上位計算機にリードデータを転送する際に、読み出しデータからＬＲＣ<Longitudinal Redundancy Check>コードを生成し、読み出しデータが最後にハードディスクに書き込まれた際にデータ転送制御部５が読み出しデータ末尾に付加したＬＲＣコードと比較する（図３のフローチャート１１５）。ディスクアレイ制御部１がＬＲＣコードの不一致を確認したら（同１１６）、ＭＰＵ３がデュアルＩ／Ｆスイッチコネクタ４０、４１を切替え（同１１７）、図５の上側に示す通信経路の接続構成から下側に示す通信経路の接続構成に変更し、同一データのリードを再実行する（同１１８）。再びデータ転送制御部５が上位計算機にリードデータを転送する際に、読み出しデータからＬＲＣコードを生成し、読み出しデータが最後にハードディスクに書き込まれた際にデータ転送制御部５が読み出しデータ末尾に付加したＬＲＣコードと比較する（同１１９）。再び上位計算機がＬＲＣコードの不一致を確認したら（同１２０）、ハードディスク１４または切替えた通信経路に故障が発生していると考えられるため（同１２１）、縮退モード（故障の認められたハードディスクを除いて、残りのハードディスクに対してリード／ライトアクセスを行う）にてディスクアレイ装置の運用を継続する（同１２２）。一方ディスクアレイ制御部１がＬＲＣコードの一致を確認したら（同１２０）、下位Ｉ／Ｆ制御部８に故障またはケーブル４９に断線が発生していると考えられるが（同１２３）、これらの故障箇所は無視されるため、ディスクアレイ装置のその後の運用には何も問題はない。したがってそのままディスクアレイ装置の運用を継続する（同１２４）。運用終了後ディスクアレイ装置の電源をいったん落とし（同１２５）、下位Ｉ／Ｆ制御部８またはケーブル４９を含む図８で示したケーブルの交換を行う（同１２６）。一方ライト時でかつライト異常が発生した場合（同１１２）、ＭＰＵが管理情報として専用のメモリに持っているデータのマッピングテーブルに、ライトの異常実行に関する情報を書き込む（同１１３）。その後ライトしたデータをリードする時点になると（同１１４）、ＭＰＵがマッピングテーブルを参照し、異常の認められたＩ／Ｆチャネルのデュアルｉ／Ｆスイッチコネクタ４０、４１を切替え、図５の上側に示す通信経路の接続構成から下側に示す通信経路の接続構成に変更し（図４のフローチャート１２７）、ライトしたと思われるデータのリードを実行する（同１２８）。読み出しデータからＬＲＣコードを生成し、読み出しデータが最後にハードディスクに書き込まれた際にデータ転送制御部５が読み出しデータ末尾に付加したＬＲＣコードと比較する（同１２９）。ディスクアレイ制御部１がＬＲＣコードの不一致を確認したら（同１３０）、故障の発生箇所はハードディスク１２と特定できるため（同１３１）、前述のリード時と同様に、縮退モードにてディスクアレイ装置の運用を継続する（同１３２）。一方、ディスクアレイ制御部１がＬＲＣコードの一致を確認したら（同１３０）、下位Ｉ／Ｆ制御部８に故障またはケーブル４９に断線が発生していると考えられるが（同１３３）、ハードディスク１２からリードするデータはライトする前の旧データと考えられるため、残りのハードディスク１３、１５、１６、１７にライトしたデータ及びパリティから書き込みに失敗したデータを再生し、上位計算機に転送すると共にハードディスク１２に再生データをライトする（同１３４）。尚前述の故障箇所は、ディスクアレイ装置のその後の運用には何も問題はないため、そのまま運用を継続する（同１３５）。運用終了後ディスクアレイ装置の電源を一度落とし（同１３６）、下位Ｉ／Ｆ制御部８またはケーブル４９を含む図８で示したケーブルの交換を行う（同１３７）。

8

【００１５】次に本発明で適用したデュアルＩ／Ｆスイッチコネクタ４０～４５の機能について説明する。

【００１６】図６は図１のデュアルＩ／Ｆスイッチコネクタ４０～４５内部の詳細を表した図である。

【００１７】デュアルＩ／Ｆスイッチコネクタ４０～４５は、図１の下位Ｉ／Ｆ制御部７～１２とコネクタにより接続し、図１のディスクアレイ制御部１がハードディスク１３～１７に対して送出するコマンド及びデータの転送経路の入口、またハードディスク１３～１７がディスクアレイ制御部１に対して送出するコマンド及びデータの転送経路の出口となる下位Ｉ／Ｆ制御部コネクタ接続ポート６０、下位Ｉ／Ｆ制御部及びケーブル断線等のハードエラーの発生によって、ディスクアレイ制御部１が送出する下位Ｉ／Ｆ制御部７～１２とハードディスク１３～１７の通信経路を切替る制御信号を受ける、スイッチ切替（以下ＳＷＣＴＬ<Switch Control>と記す）信号入力ポート６１、下位Ｉ／Ｆ制御部７～１２とハードディスク１３～１７の間の通信方向を切替る制御信号を受ける、通信方向切替（以下ＷＲＥＮ<Write Enable>と記す）信号入力ポート６２、ディスクアレイ装置の正常運用時に使用する主経路の有効／無効及び通信方向を切替る手段６３、ディスクアレイ装置の下位Ｉ／Ｆ制御部７～１２及びケーブル４６～５５の断線等のハードエラー時に使用する、予備経路の有効／無効及び通信方向を切替る手段６４、ＳＷＣＴＬ信号に従って、予備経路の有効／無効及び通信方向を切替る手段６４のＷＲＥＮ信号の入力の有効／無効を切替る手段６５、ハードディスク１３～１７とコネクタにより接続し、ハードディスク１３～１７がディスクアレイ制御部１に対して送出するコマンド及びデータの転送経路の入口、また下位Ｉ／Ｆ制御部７～１２がハードディスク１３～１７に対して送出するコマンド及びデータの転送経路の出口となる

9 10

ハードディスクコネクタ接続ポート６６、６７で構成される。

【００１８】

【表１】

表１　デュアルＩ／Ｆスイッチコネクタ制御信号仕様一覧

		SWCTL−N	WREN−N
MAIN	Write	偽	真
	Read		偽
SUB	Write	真	真
	Read		偽

【００１９】表１に、デュアルＩ／Ｆスイッチコネクタ制御信号の一覧を示す。

【００２０】図７のフローチャートに基づき本発明のディスクアレイ装置のデュアルＩ／Ｆスイッチコネクタの内部動作について説明する。

【００２１】ディスクアレイ装置の電源を入れると（図７のフローチャート２００）、図１のディスクアレイ制御部１のＭＰＵ３がデュアルＩ／Ｆスイッチコネクタ４０～４５のＳＷＣＴＬ信号を偽に設定する（同２０１）。したがって主経路の通信経路を用いてディスクアレイ装置の運用を開始する（同２０２）。まず上位計算機がリード／ライトコマンドをディスクアレイ制御部１に送出すると（同２０３）、リードコマンドを受けた場合（同２０４）、ディスクアレイ制御部１のＭＰＵ３がデュアルＩ／Ｆスイッチコネクタ４０～４５のＷＲＥＮ信号を偽に設定する（同２０５）。一方ライトコマンドを受けた場合は（同２０４）、ＷＲＥＮ信号を真に設定する（同２０６）。以上の設定の後、ハードディスク１３～１７に対してリード／ライトアクセスを実行する（同２０７）。いずれの下位Ｉ／Ｆ制御部７～１２、ケーブル４６～５５及びハードディスク１３～１７に故障が発生していなければ、前述の設定は変えずにリード／ライトアクセスを継続し、アクセス終了後ディスクアレイ制御部１が上位計算機から次のリード／ライトコマンドの要求を受けなければ（同２１１）、ディスクアレイ装置の運用は終了する（同２１２）。一方次のリード／ライトコマンドの要求を受ければ（同２１１）、引続き次のリード／ライトアクセスを実行する。しかし、リード／ライトアクセス実行中に下位Ｉ／Ｆ制御部７～１２、ケーブル４６～５５及びハードディスク１３～１７のいずれかに故障が発生した場合には（同２０８）、ディスクアレイ制御部１のＭＰＵ３がデュアルＩ／Ｆスイッチコネクタ４０～４５のＳＷＣＴＬ信号を真に設定する（同２０９）。したがって予備経路を用いてディスクアレイ装置の運用を継続する。以後ディスクアレイ装置の運用を終了するまで、前述の設定は変更しない。

【００２２】このように本方式のディスクアレイ装置は、従来の下位Ｉ／Ｆ制御部とハードディスクを一対一対応に接続している従来のディスクアレイ装置に比較して、予備の下位Ｉ／Ｆ制御部を付加することで、１台の下位Ｉ／Ｆ制御部の故障発生もしくは１本のケーブルの断線がディスクアレイ装置の運用に全く影響を及ぼさないという特徴を持っている。これはつまり本方式を適用することで、ディスクアレイ装置の下位Ｉ／Ｆ制御部、ハードディスク及びそれらを相互に接続するケーブルを含む下位Ｉ／Ｆチャネルに対する耐故障の信頼性を高めることができることを意味している。また図１の下位Ｉ／Ｆ制御部７～１２にソケット式のものを用いることで、ディスクアレイ制御部１を構成する基盤を交換せずに故障の発生した下位Ｉ／Ｆ制御部のみを交換することで、修理コストを低減できる。さらにデュアルＩ／Ｆスイッチコネクタ４０～４５とハードディスク１３～１７の間に使用するケーブルの形状も、２本ケーブルの一端を一つに束ねていることを除けば従来のケーブルと同等であるため、取り扱いは用意である。

【００２３】第一実施例ではディスクアレイ制御部に対して新たな予備の下位Ｉ／Ｆ制御部を付加する方法を提示したが、特にハードウエアのコストを抑えてディスクアレイ装置を構成する必要がある場合には、ディスクアレイ制御部にすでに搭載している本来バックアップ機器を接続するために使用する下位Ｉ／Ｆ制御部を予備の下位Ｉ／Ｆ制御部として流用する方法が考えられる。この方法について次の第二実施例で述べる。

【００２４】本発明の第二実施例を図９、図１０を用いて説明する。

【００２５】図９は従来のディスクアレイ装置が外部機器の接続用に持っている下位Ｉ／Ｆ制御部を、第一実施例で述べた本発明のディスクアレイ装置に付加した予備の下位Ｉ／Ｆ制御部に流用したディスクアレイ装置全体を表した図である。ディスクアレイ制御部１のハードウエア構成は第一実施例で述べたディスクアレイ装置と同様であるが、図６で示したデュアルＩ／Ｆスイッチコネクタ４０、４５の各々の二つのハードディスクコネクタ接続ポートの内１つが空いているため、ここにバックアップ機器用Ｉ／Ｆコネクタ５６、５７を設ける。ここではバックアップ機器用Ｉ／Ｆコネクタ５６にバックアップ機器を接続して、ハードディスク１３～１７のデータをバックアップする方法について説明する。まずディスクアレイ装置がリード／ライトアクセス実行中に（図１０のフローチャート３００）上位計算機がハードディスク１３～１７のデータのバックアップを要求すると（同３０１）、データ転送制御部５がハードディスク１３～１７からデータを収集し一時データバッファ４へ転送保持する（同３０２）。下位Ｉ／Ｆ制御部、使用中のケー

ブルに故障が発生していない間は、主経路であるケーブル４７、４９、５１、５３、５５がディスクアレイ制御部１とハードディスク１３～１７間の通信経路になるが、バックアップ機器はバックアップ機器用Ｉ／Ｆコネクタ５６を介してデュアルＩ／Ｆスイッチコネクタ４０の主経路を利用できるため、下位Ｉ／Ｆ制御部７をバックアップ機器専用の下位Ｉ／Ｆ制御部として使うことができる。したがって、データバッファ４に保持しているバックアップデータを即バックアップ機器に転送する（同３０９）。バックアップ完了後（同３１０）、再びハードディスク１３～１７に対するリード／ライトアクセスを継続する。一方いずれの下位Ｉ／Ｆ制御部、使用中のケーブルに故障が発生した場合には（同３０３）、デュアルＩ／Ｆスイッチコネクタの予備経路とケーブル４６がハードディスク１３との通信経路になるため、ハードディスク１３に対するリード／ライトアクセスを優先する。上位計算機がハードディスク１３～１７のデータのバックアップを要求すると（同３０１）、データ転送制御部５がハードディスク１３～１７からデータを収集し一時データバッファ４へ転送保持する（同３０２）。ハードディスク１３に対する次のリード／ライトアクセス要求が来ない間に（同３０４）、デュアルＩ／Ｆスイッチコネクタ４０を主経路に切替え（同３０５）、データバッファ４に保持しているバックアップデータをバックアップ機器に転送する（同３０６）。バックアップ完了後（同３０７）、デュアルＩ／Ｆスイッチコネクタ４０を予備経路に切替え（同３０８）、再びハードディスク１３～１７に対してリード／ライトアクセスを継続する。

【００２６】第二実施例で示したバックアップ方法を利用すると、正常時のみならずリード／ライトアクセス中に下位Ｉ／Ｆ制御部８～１２、ケーブル４７、４９、５１、５３、５５に故障が発生した場合にもハードディスク１３～１７のデータのバックアップができるため、データ喪失の発生を抑えることができる。しかし、バックアップに使用中の下位Ｉ／Ｆ制御部７、バックアップ機器用Ｉ／Ｆコネクタ５６が故障する可能性も考えられるため、この問題に対処するためにハードディスク１３～１７から転送したバックアップデータをバッファメモリ４に保持する時にバックアップデータのコピーをつくり、バックアップ処理が完了するまでそのコピーデータをバッファメモリ４に保持しておく。したがって、バックアップデータをバックアップ機器１９に転送中に前述の故障が発生してもコピーデータはバッファメモリ４に残っているため、バックアップ機器１９をバックアップ機器用Ｉ／Ｆコネクタ５７につなぎ替えることでバックアップを再実行することができる。この場合もハードディスク１３～１７に対するリード／ライトアクセスを優先するため、ハードディスク１７に対する次のリード／ライトアクセス要求が来ない間にデュアルＩ／Ｆスイッチコネクタ４５を主経路から予備経路に切替えバックアップを実行し、バックアップ完了後デュアルＩ／Ｆスイッチコネクタ４５を予備経路から主経路に戻し、再びハードディスク１３～１７に対してリード／ライトアクセスを継続する。ここで挙げた方式は、バッファメモリ４にバックアップデータのコピーを保持できる空き容量のある場合には特に勧められるが、空き容量に余裕がない場合も考えられる。この問題を考慮したバックアップ方法を次の第三実施例で述べる。

【００２７】本発明の第三実施例を図１１を用いて説明する。

【００２８】図１１は第二実施例で述べたディスクアレイ装置を変形させたディスクアレイ装置全体を表した図である。ディスクアレイ制御部１のハードウエア構成は第二実施例で述べたディスクアレイ装置と同様であるが、デュアルＩ／Ｆスイッチコネクタ４０、４５の各々１つの空いているハードディスクコネクタ接続ポートにそれぞれ１本のケーブルを接続し、それら２本のケーブルの終端にＭＰＵで操作するバックアップデータ転送経路切替スイッチ５８を設ける。そしてバックアップデータ転送経路切替スイッチ５８から出ているバックアップ機器用Ｉ／Ｆコネクタ５９にバックアップ機器１９を接続する。ここで述べるディスクアレイ装置のバックアップ動作は第二実施例で示したディスクアレイ装置とほぼ同様であるが、バックアップデータのコピーを作りバッファメモリ４に保持しておく必要はない。それはつまり、下位Ｉ／Ｆ制御部７でバックアップを実行中に、下位Ｉ／Ｆ制御部７及びそれとバックアップデータ転送経路切替スイッチ５８をつないでいるケーブルに故障が発生しても、ＭＰＵがバックアップデータ転送経路切替スイッチ５８を図１１に示すポジションＣからＡに切替えることで、下位Ｉ／Ｆ制御部１２を用いてバックアップデータ転送を継続することができるからである。また上位計算機からバックアップの要求を受けていない間は、ＭＰＵがバックアップデータ転送経路切替スイッチ５８を図１１に示すポジションＢに設定するので、バックアップ機器１９の取り外しや交換はディスクアレイ装置の電源を落とさずに自由にできる。本ディスクアレイ装置は、バックアップデータ転送経路切替スイッチ５８と僅かなケーブルを付加することで、いずれの下位Ｉ／Ｆチャネルの故障発生に関わらず、正常運用時と同様にハードディスクのデータをバックアップできる特徴を持っている。

【００２９】特に第二、第三実施例で述べたディスクアレイ装置は、予備の下位Ｉ／Ｆ制御部に本来外部機器専用の下位Ｉ／Ｆ制御部を流用することから、従来のディスクアレイ装置とほとんど変わらないハードウエアコストで、下位Ｉ／Ｆチャネルの耐故障性を高めことができる方法として勧められる。

【００３０】ところで、従来のディスクアレイ装置のハ

ードウエア構成を変えずに、拡張ボード（以下ドータボードと記す）を付加することで、第一から第三実施例で提案したのと同等な下位Ｉ／Ｆチャネルの耐故障性の高いディスクアレイ装置を実現する方法も考えられる。この方法について次の第四実施例で述べる。

【００３１】本発明の第四実施例を図１２を用いて説明する。

【００３２】図１２はデュアルＩ／Ｆスイッチコネクタ及びバックアップデータ転送経路切替えスイッチを搭載したドータボードを付加したディスクアレイ装置全体図である。ディスクアレイ制御部１を構成する要素は従来のディスクアレイ装置と全く同様であり、第一から第三実施例までディスクアレイ制御部１を構成する基盤の内部に搭載していたデュアルＩ／Ｆスイッチコネクタ４０～４５及びバックアップデータ転送経路切替えスイッチ５８を一枚のドータボード６８に搭載し、ディスクアレイ制御部１とハードディスク１３～１７の間に設置し、図８で示した形状のケーブルでデュアルＩ／Ｆスイッチコネクタ４０～４５とハードディスク１３～１７を接続することで、従来のディスクアレイ装置を下位Ｉ／Ｆチャネルの耐故障性の高いディスクアレイ装置へ容易にアップグレードできることを示している。

【００３３】第一から第四実施例は下位Ｉ／Ｆ制御部にＳＣＳＩを用いたディスクアレイ装置を元にしているが、下位Ｉ／Ｆチャネルの耐故障性の高いディスクアレイ装置を提供する本発明は、次世代のＩ／Ｆ制御部として注目されているシリアルＩ／Ｆを用いたディスクアレイ装置に対しても適用できるものである。本発明を適用したシリアルＩ／Ｆを用いたディスクアレイ装置について次の第五実施例で述べる。

【００３４】第五実施例を図１３を用いて説明する。

【００３５】図１３はシリアルＩ／Ｆを用いた本発明のディスクアレイ装置全体図である。ディスクアレイ制御部６９には、図１７に示したシリアルＩ／Ｆを用いた従来のディスクアレイ装置のディスクアレイ制御部２９に、シリアルＩ／Ｆ制御部２２、シリアルＩ／Ｆ制御部２２を制御する専用のシリアルＩ／Ｆ制御用ＭＰＵ２３と同等の機能を持つ予備用Ｉ／Ｆ制御部として、シリアルＩ／Ｆ制御部７０、シリアルＩ／Ｆ制御部７０を制御する専用のシリアルＩ／Ｆ制御用ＭＰＵ７１を持つ。またデータ転送制御部５とシリアルＩ／Ｆ制御部２２またはシリアルＩ／Ｆ制御部７０に接続を切替える通信経路切替えスイッチ７２、さらにシリアルＩ／Ｆ制御部２２とハードディスク２４～２８間及びシリアルＩ／Ｆ制御部とハードディスク２４～２８間の通信経路の有効／無効を切替えるデュアルＩ／Ｆスイッチコネクタ７３をディスクアレイ制御部６９内部に持つ。本発明のディスクアレイ装置は予備用Ｉ／Ｆ制御部を備えることで、リード／ライトアクセス実行中にシリアルＩ／Ｆ制御部２２に故障が発生した場合にも、リード／ライトアクセスの

異常実行によって発生するデータの不一致を基に、ＭＰＵ３が通信経路切替えスイッチ７２及びデュアルＩ／Ｆスイッチコネクタ７３を予備のシリアルＩ／Ｆ制御部７０が使用できるように切替え、実行中のリード／ライトアクセスを継続することができる。尚ディスクアレイ装置の運用が終了したら、一度ディスクアレイ装置の電源を落とし、ディスクアレイ制御部６９を構成するメインボードを交換するか、もしくはソケット式のシリアルＩ／Ｆ制御部を用いていれば、シリアルＩ／Ｆ制御部のみを交換する。

【００３６】次にシリアルＩ／Ｆを用いた本発明のディスクアレイ装置に適用したデュアルＩ／Ｆスイッチコネクタ７３の機能について説明する。

【００３７】図１４は図１３のデュアルＩ／Ｆスイッチコネクタ７３内部の詳細を表した図である。デュアルＩ／Ｆスイッチコネクタ７３は、図１３のシリアルＩ／Ｆ制御部２２、７０とコネクタにより接続し、図１３のディスクアレイ制御部６９がハードディスク２４～２８に対して送出するコマンド及びデータの転送経路の入口、またハードディスク２４～２８がディスクアレイ制御部６９に対して送出するコマンド及びデータの転送経路の出口となるディスクアレイ制御部側シリアルＩ／Ｆ制御部通信入出力ポート７５～８２、シリアルＩ／Ｆ制御部２２のハードエラーの発生時に、シリアルＩ／Ｆ制御部２２とハードディスク２４～２８から、シリアルＩ／Ｆ制御部７０とハードディスク２４～２８の通信経路に切替るため、ディスクアレイ制御部６９が送出する制御信号を受けるスイッチ切替（以下ＳＷＣＴＬと記す）信号入力ポート７４、ディスクアレイ装置の正常運用時に使用する主経路の有効／無効及び通信方向を切替る手段８３～８６、ディスクアレイ装置のシリアルＩ／Ｆ制御部２２のハードエラー時に使用する、予備経路の有効／無効及び通信方向を切替る手段８７～９０、ハードディスク２４～２８とコネクタにより接続し、ハードディスク２４～２８がディスクアレイ制御部６９に対して送出するコマンド及びデータの転送経路の入口、またシリアルＩ／Ｆ制御部２２またはシリアルＩ／Ｆ制御部７０がハードディスク２４～２８に対して送出するコマンド及びデータの転送経路の出口となるハードディスクコネクタ側シリアルＩ／Ｆ制御部通信入出力ポート９１～９４で構成される。ディスクアレイ装置の電源投入後（図１５のフローチャート４００）、ディスクアレイ制御部６９のＭＰＵ３がＳＷＣＴＬ信号を偽に設定する（同４０１）。したがって主経路の有効／無効及び通信方向を切替る手段８３～８６は有効になり、予備経路の有効／無効及び通信方向を切替る手段８７～９０は無効になる（同４０２）。通信経路の確定後上位計算機がディスクアレイ制御部６９にリード／ライトコマンドを送出し（同４０３）、ハードディスク２４～２８に対するリード／ライトアクセスを開始する（同４０４）。その後リ

ード／ライトアクセス中にシリアルＩ／Ｆ制御部２２に故障が発生しなければ（同４０５）、現在実行中のリード／ライトアクセス終了後（同４０６）、上位計算機が次のコマンドを送出している場合、引続きリード／ライトアクセスを実行し（同４０７）、コマンド要求がなければディスクアレイ装置の運用を終了する（同４０８）。したがって正常運用時は主経路のみを用いることになる。一方リード／ライトアクセス中にシリアルＩ／Ｆ制御部２２に故障が発生すると（同４０５）、ディスクアレイ制御部６９のＭＰＵ３がＳＷＣＴＬ信号を真に設定する（同４０９）。したがって主経路の有効／無効及び通信方向を切替る手段８３～８６は無効になり、予備経路の有効／無効及び通信方向を切替る手段８７～９０は有効になる（同４１０）。尚その後に上位計算機が次のコマンドを送出している場合、引続きリード／ライトアクセスを実行するが（同４０７）、ディスクアレイ装置の運用を終了するまで前述の設定は変更しない。

【００３８】第五実施例で述べた本発明のディスクアレイ装置は、ディスクアレイ制御部の下位Ｉ／Ｆ制御部にシリアルＩ／Ｆを用いた従来のディスクアレイ装置の信頼性の問題を補う方式として提案できるものである。つまり各々のハードディスクの故障には２組ある送信受信ポートでディスクアレイ制御部との通信経路を常に確保できるが、コマンド及びデータをハードディスクに対して送信受信するディスクアレイ制御部本体の送信受信ポートでは故障を回避することが出来ない。

【００３９】

【発明の効果】本発明によれば、下位Ｉ／Ｆ制御部、ケーブルの故障によるディスクアレイ装置の縮退モードによる運用（故障したＩ／Ｆチャネルのハードディスクを除いた、残りのハードディスクによるディスクアレイ装置の運用）が発生しないため、特にリードアクセス中にデータ転送性能の劣化のない連続運用が提供できる。

【００４０】従来のディスクアレイ装置に下位Ｉ／Ｆチャネルの故障を回避する回路を搭載したドータボードを付加することで、装置の信頼性を高めるアップグレード方法を提供できる。

【００４１】従来の下位Ｉ／Ｆ制御部にシリアルＩ／Ｆ制御部を用いたディスクアレイ装置に、下位Ｉ／Ｆチャネルの耐故障性を高める方法を提供できる。

【図面の簡単な説明】

【図１】本発明のディスクアレイ装置のブロック図。

【図２】本発明のディスクアレイ装置の動作を表したフローチャート。

【図３】本発明のディスクアレイ装置の動作を表したフローチャート。

【図４】本発明のディスクアレイ装置の動作を表したフローチャート。

【図５】実施例１における本発明のディスクアレイ装置の下位通信経路接続構成の説明図。

【図６】本発明のディスクアレイ装置に適用したデュアルＩ／Ｆスイッチコネクタの詳細を表した説明図。

【図７】デュアルＩ／Ｆスイッチコネクタの内部動作を表したフローチャート。

【図８】本発明のディスクアレイ装置に適用したハードディスクとデュアルＩ／Ｆスイッチコネクタを接続するケーブルの外観を表した説明図。

【図９】本発明のディスクアレイ装置にバックアップ装置を付加した状態を表した説明図。

【図１０】バックアップ装置を付加した本発明のディスクアレイ装置のバックアップ動作を表したフローチャート。

【図１１】本発明のディスクアレイ装置にバックアップ装置を付加した状態を表した説明図。

【図１２】従来のディスクアレイ装置にデュアルＩ／Ｆスイッチコネクタを搭載したドータボードを適用した状態を表した説明図。

【図１３】本発明のディスクアレイ装置（シリアルＩ／Ｆ使用）の説明図。

【図１４】本発明のディスクアレイ装置（シリアルＩ／Ｆ使用）に適用したデュアルＩ／Ｆスイッチコネクタの詳細を表した説明図。

【図１５】デュアルＩ／Ｆスイッチコネクタ（シリアルＩ／Ｆ使用）の内部動作を表したフローチャート。

【図１６】従来のディスクアレイ装置の説明図。

【図１７】パリティ生成及び喪失データ再生の仕組の説明図。

【図１８】従来のディスクアレイ装置（シリアルＩ／Ｆ使用）の説明図。

【図１９】シリアルＩ／Ｆ（ＳＳＡＩ／Ｆ）の詳細な構成を表した説明図。

【図２０】従来のディスクアレイ装置（シリアルＩ／Ｆ使用）のハードディスクの故障時のアクセス状態を表した説明図。

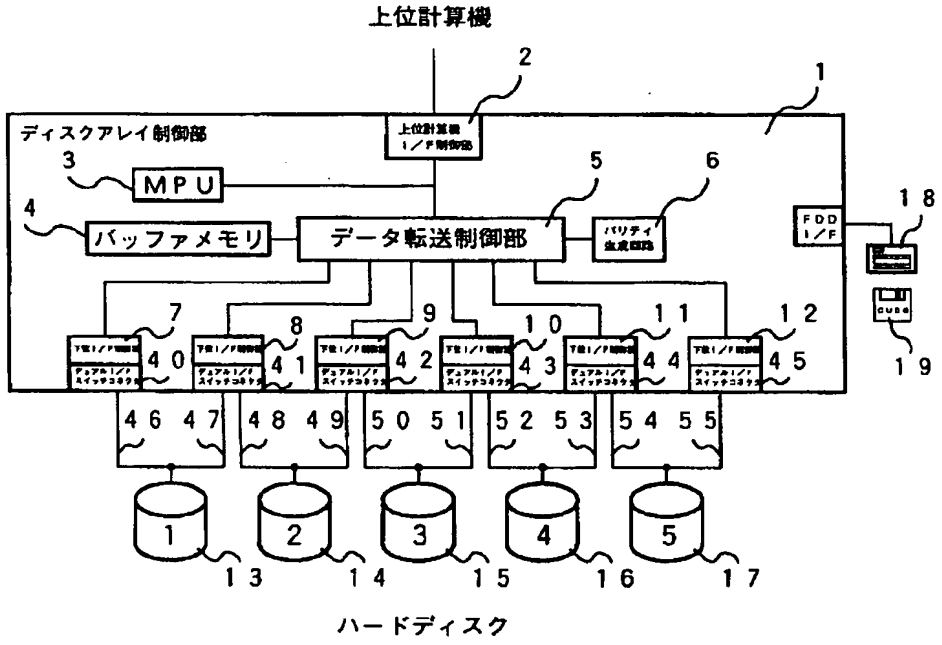
【図２１】従来のディスクアレイ装置（シリアルＩ／Ｆ使用）の複数データ転送実行状態を表した説明図。

【符号の説明】

１…ディスクアレイ制御部、
２…上位計算機Ｉ／Ｆ制御部、
３…ＭＰＵ、
４…バッファメモリ、
５…データ転送制御部、
６…パリティ生成回路、
７～１２…下位Ｉ／Ｆ制御部、
１３～１７…ハードディスク、
１８…フロッピディスクＩ／Ｆ制御部、
１９…フロッピディスクドライブ、
４０～４５…デュアルＩ／Ｆスイッチコネクタ、
４６～５５…ケーブル。

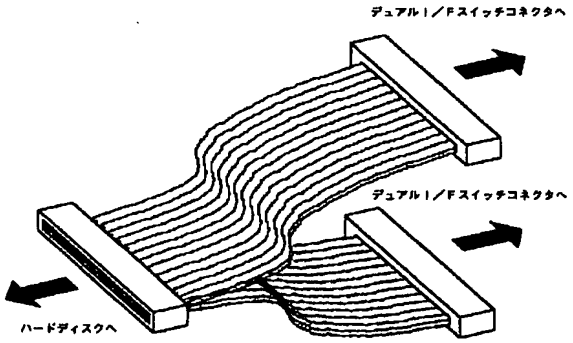
【図１】

図 1



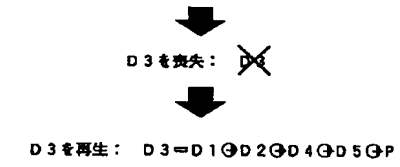
【図８】

図 8



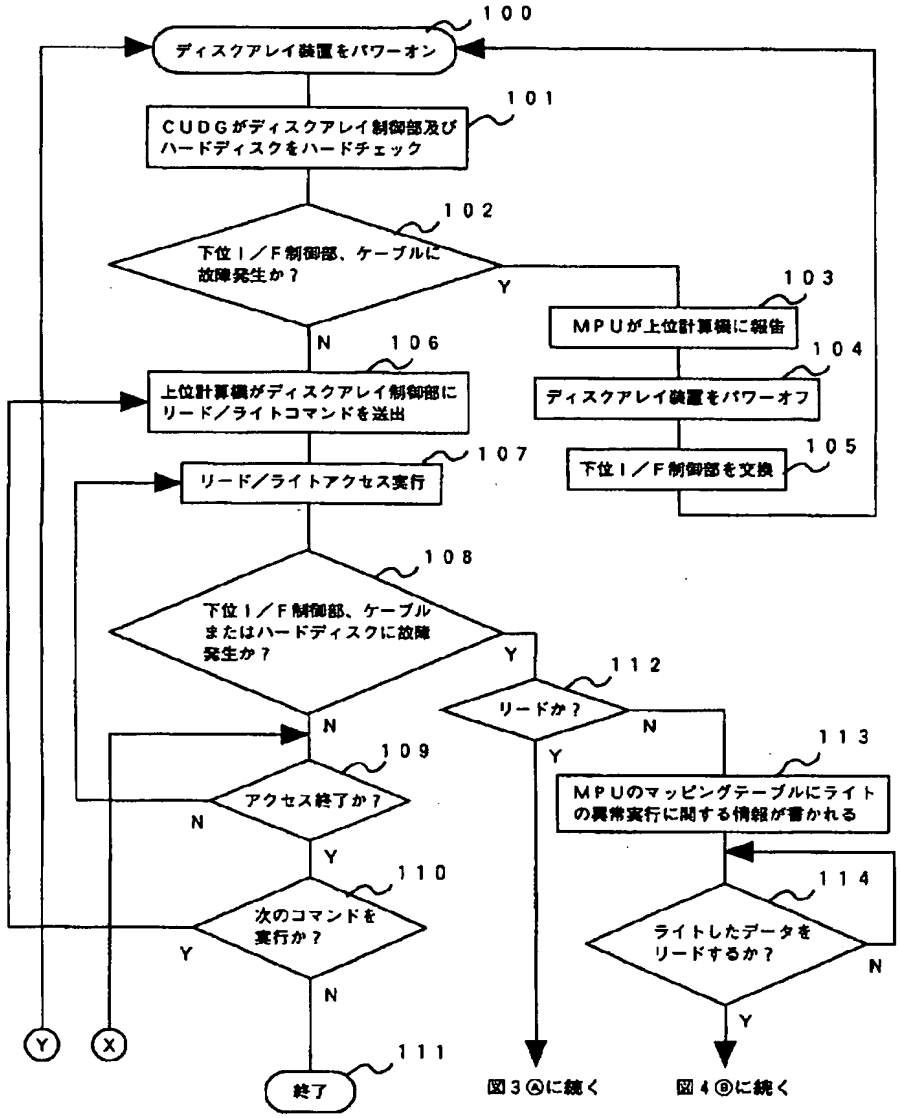
【図１７】

図 17



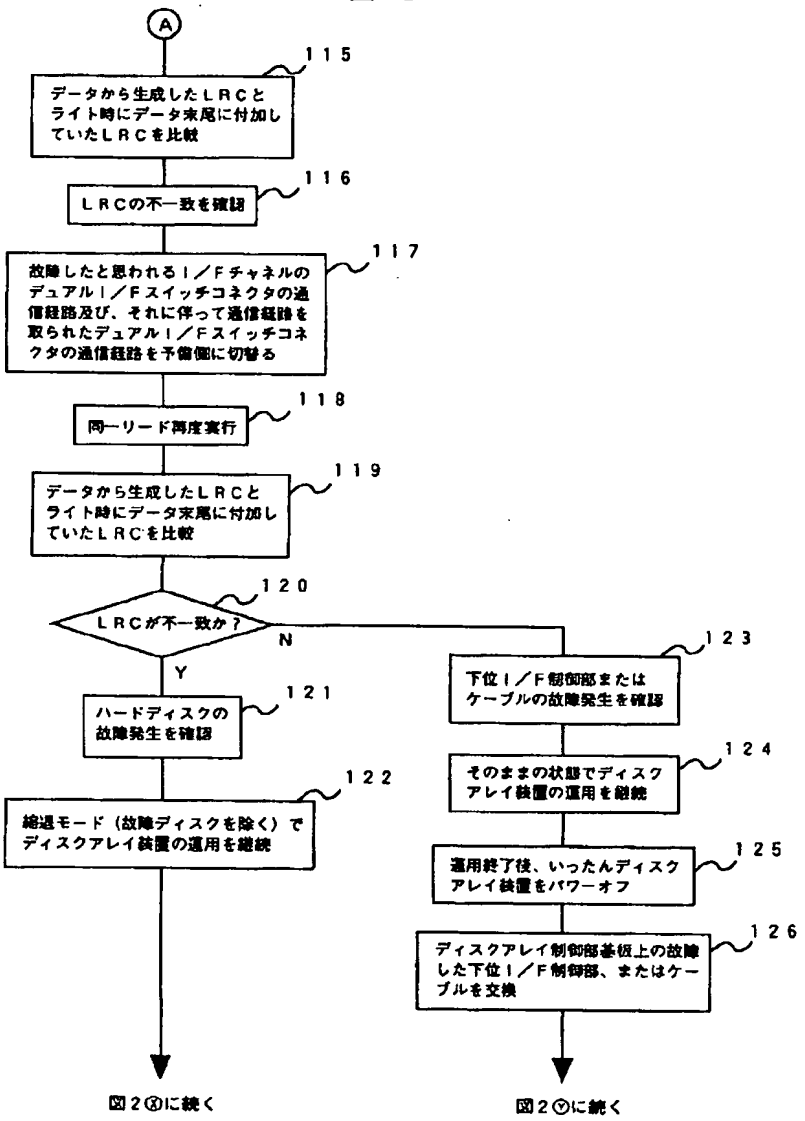
【図２】

図 2


100
ディスクアレイ装置をパワーオン
101
CUDGがディスクアレイ制御部及びハードディスクをハードチェック
102
下位Ｉ／Ｆ制御部、ケーブルに故障発生か？
Y
103
ＭＰＵが上位計算機に報告
104
ディスクアレイ装置をパワーオフ
105
下位Ｉ／Ｆ制御部を交換
N
106
上位計算機がディスクアレイ制御部にリード／ライトコマンドを送出
107
リード／ライトアクセス実行
108
下位Ｉ／Ｆ制御部、ケーブルまたはハードディスクに故障発生か？
Y
112
リードか？
N
Y
113
ＭＰＵのマッピングテーブルにライトの異常実行に関する情報が書かれる
N
109
アクセス終了か？
N
Y
110
次のコマンドを実行か？
Y
N
114
ライトしたデータをリードするか？
N
Y
Y
X
111
終了
図３Ⓐに続く
図４Ⓑに続く

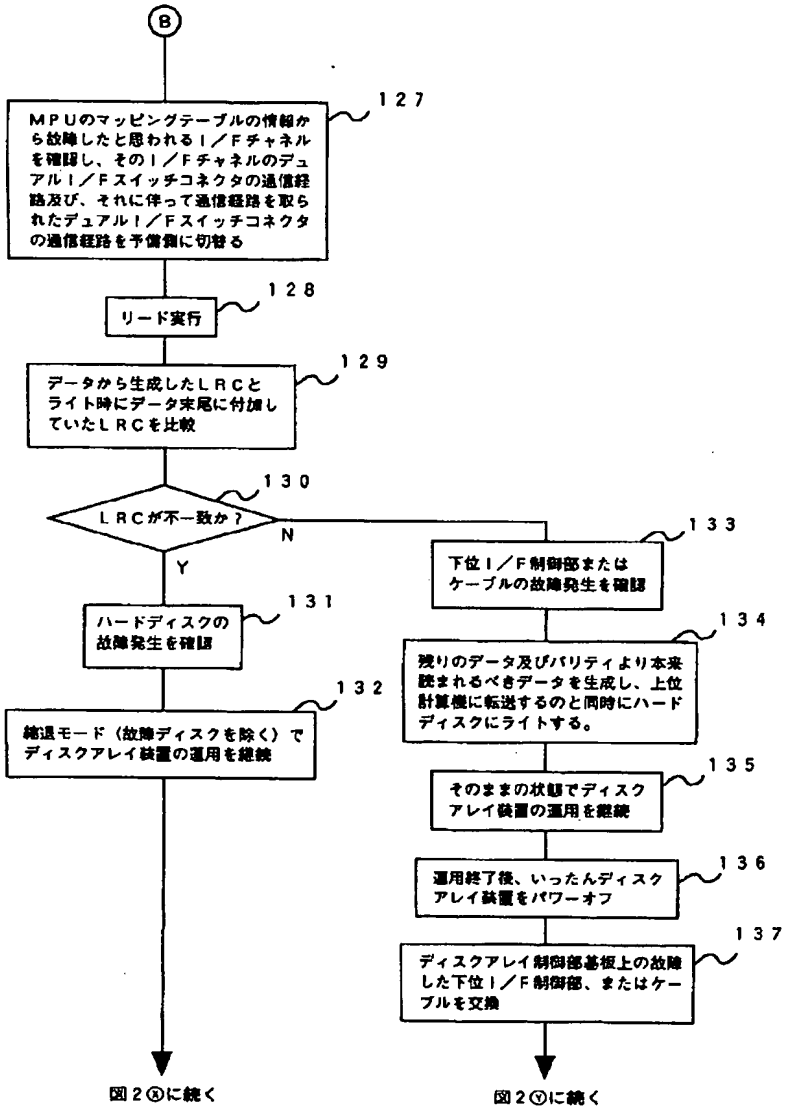
【図３】

図 3


A
115
データから生成したＬＲＣとライト時にデータ末尾に付加していたＬＲＣを比較
116
ＬＲＣの不一致を確認
117
故障したと思われるＩ／Ｆチャネルのデュアル！／Ｆスイッチコネクタの通信経路及び、それに伴って通信経路を取られたデュアル！／Ｆスイッチコネクタの通信経路を予備側に切替る
118
同一リード再度実行
119
データから生成したＬＲＣとライト時にデータ末尾に付加していたＬＲＣを比較
120
ＬＲＣが不一致か？
N
Y
121
ハードディスクの故障発生を確認
122
縮退モード（故障ディスクを除く）でディスクアレイ装置の運用を継続
123
下位！／Ｆ制御部またはケーブルの故障発生を確認
124
そのままの状態でディスクアレイ装置の運用を継続
125
運用終了後、いったんディスクアレイ装置をパワーオフ
126
ディスクアレイ制御部基板上の故障した下位！／Ｆ制御部、またはケーブルを交換
図２Ⓧに続く
図２Ⓨに続く

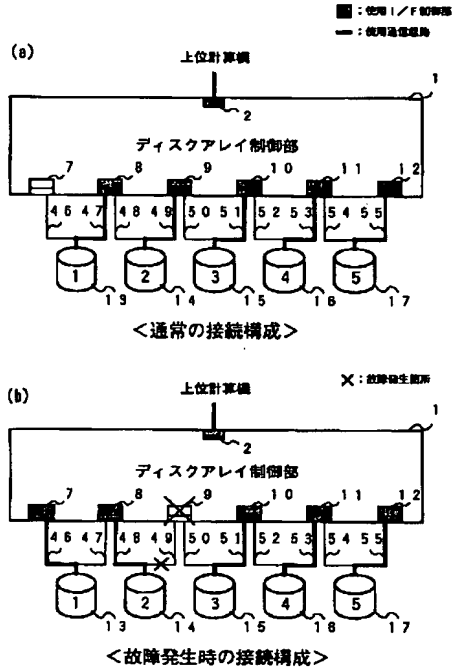
【図４】

図 4


B
MPUのマッピングテーブルの情報から故障したと思われるI／Fチャネルを確認し、そのI／Fチャネルのデュアル I／Fスイッチコネクタの通信経路及び、それに伴って通信経路を取られたデュアル I／Fスイッチコネクタの通信経路を予備側に切替る
127
リード実行
128
データから生成したLRCとライト時にデータ末尾に付加していたLRCを比較
129
LRCが不一致か？
130
N
Y
下位I／F制御部またはケーブルの故障発生を確認
133
ハードディスクの故障発生を確認
131
残りのデータ及びパリティより本来読まれるべきデータを生成し、上位計算機に転送するのと同時にハードディスクにライトする。
134
縮退モード（故障ディスクを除く）でディスクアレイ装置の運用を継続
132
そのままの状態でディスクアレイ装置の運用を継続
135
運用終了後、いったんディスクアレイ装置をパワーオフ
136
ディスクアレイ制御部基板上の故障した下位I／F制御部、またはケーブルを交換
137
図２Ⓧに続く
図２Ⓨに続く

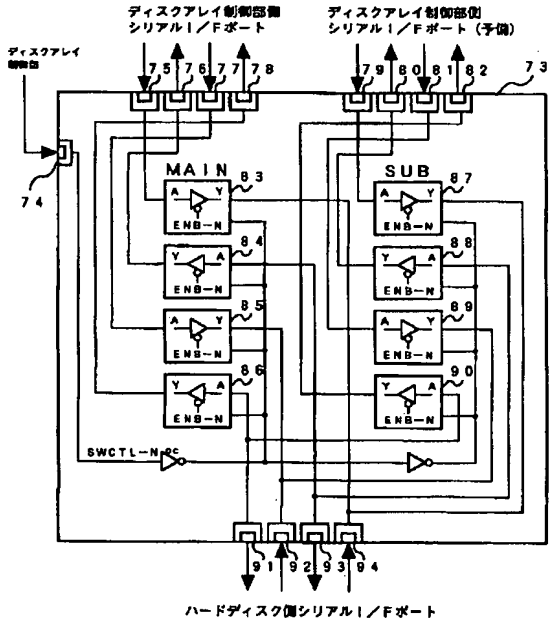
【図５】

図 5



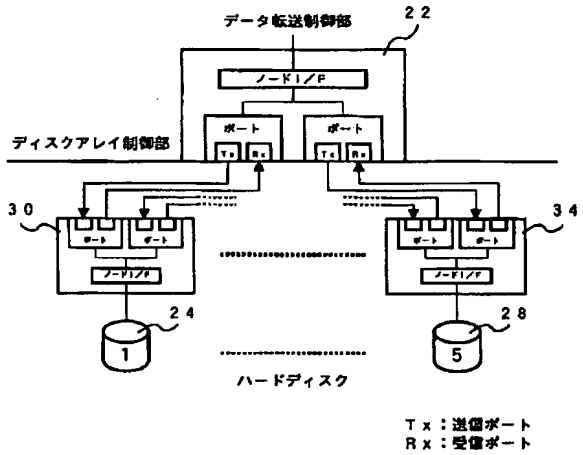
【図１４】

図 14



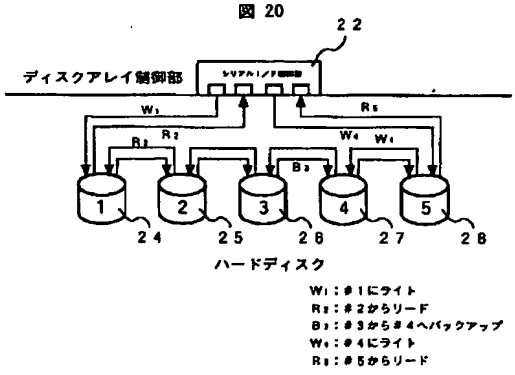
【図１９】

図 19



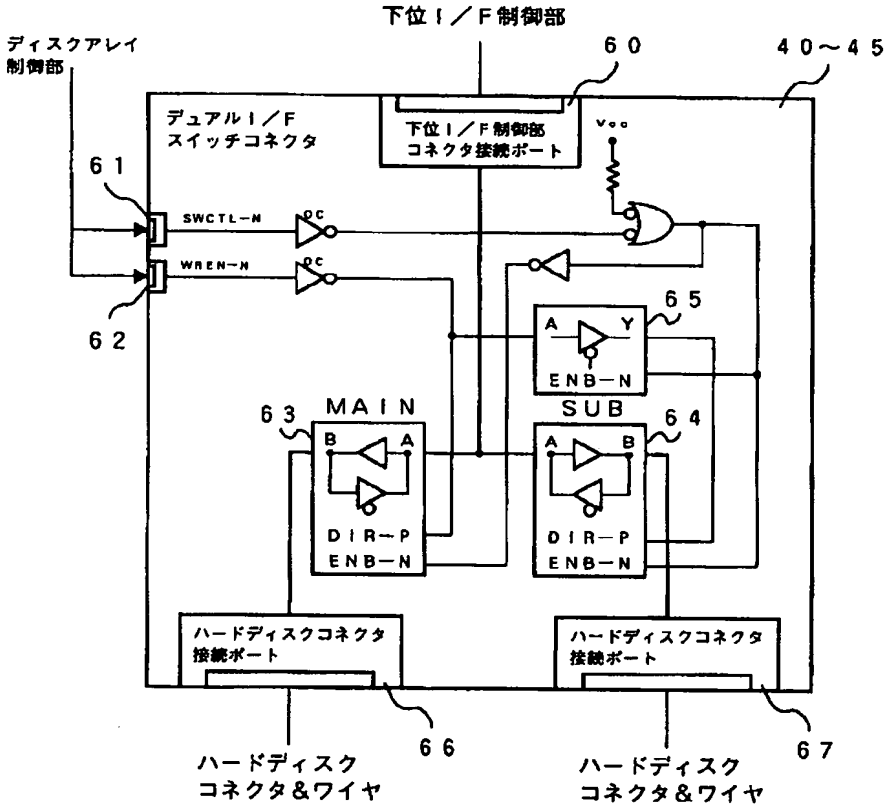
【図２０】

図 20



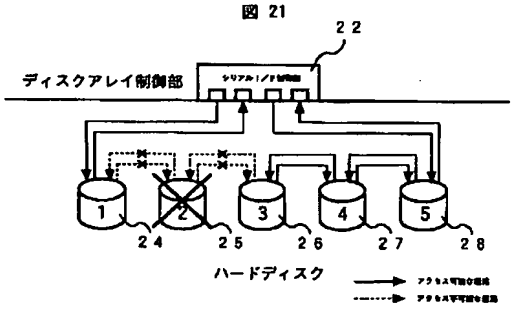
【図６】

図 6


下位Ｉ／Ｆ制御部
ディスクアレイ制御部
６０
４０～４５
デュアルＩ／Ｆスイッチコネクタ
下位Ｉ／Ｆ制御部コネクタ接続ポート
Vcc
６１
SWCTL－N
WREN－N
OC
６２
６５
A
Y
ENB－N
６３
MAIN
SUB
６４
B
A
DIR－P
ENB－N
ハードディスクコネクタ接続ポート
６６
６７
ハードディスクコネクタ＆ワイヤ

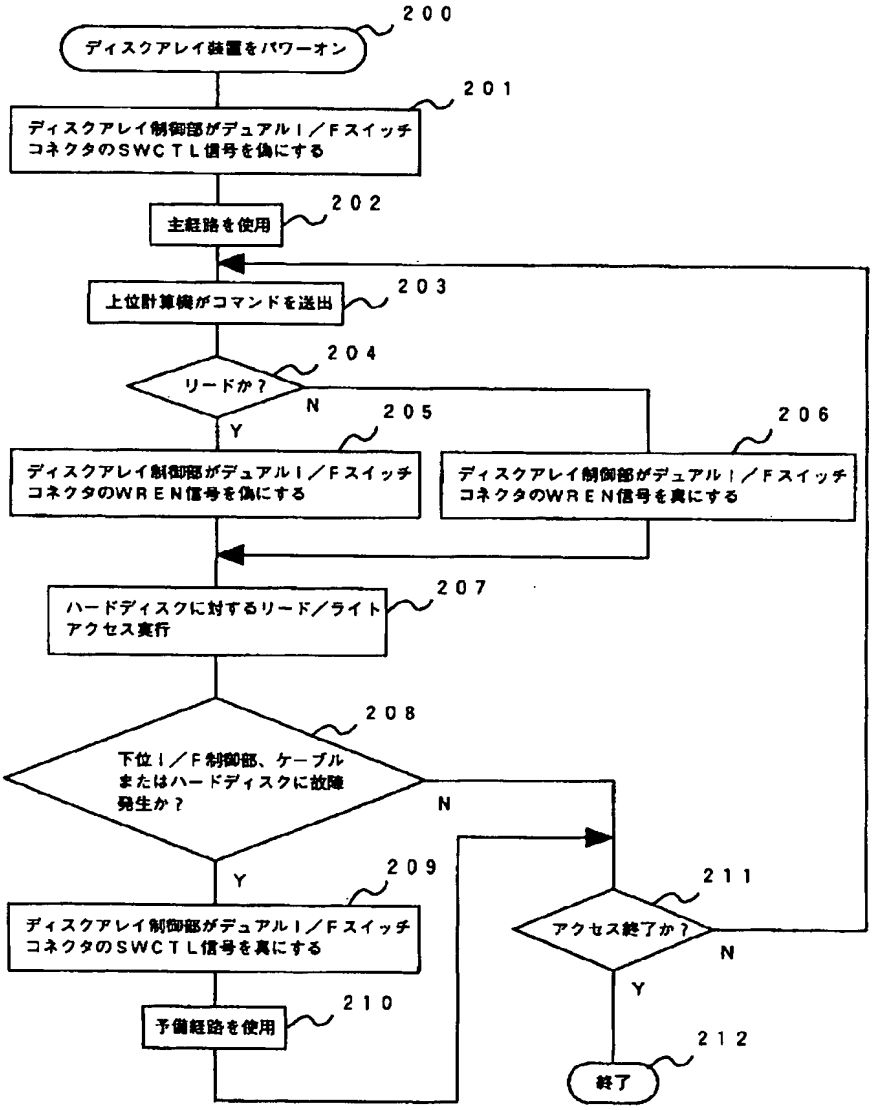
【図２１】

図 21


２２
ディスクアレイ制御部
1
2
3
4
5
２４
２５
２６
２７
２８
ハードディスク

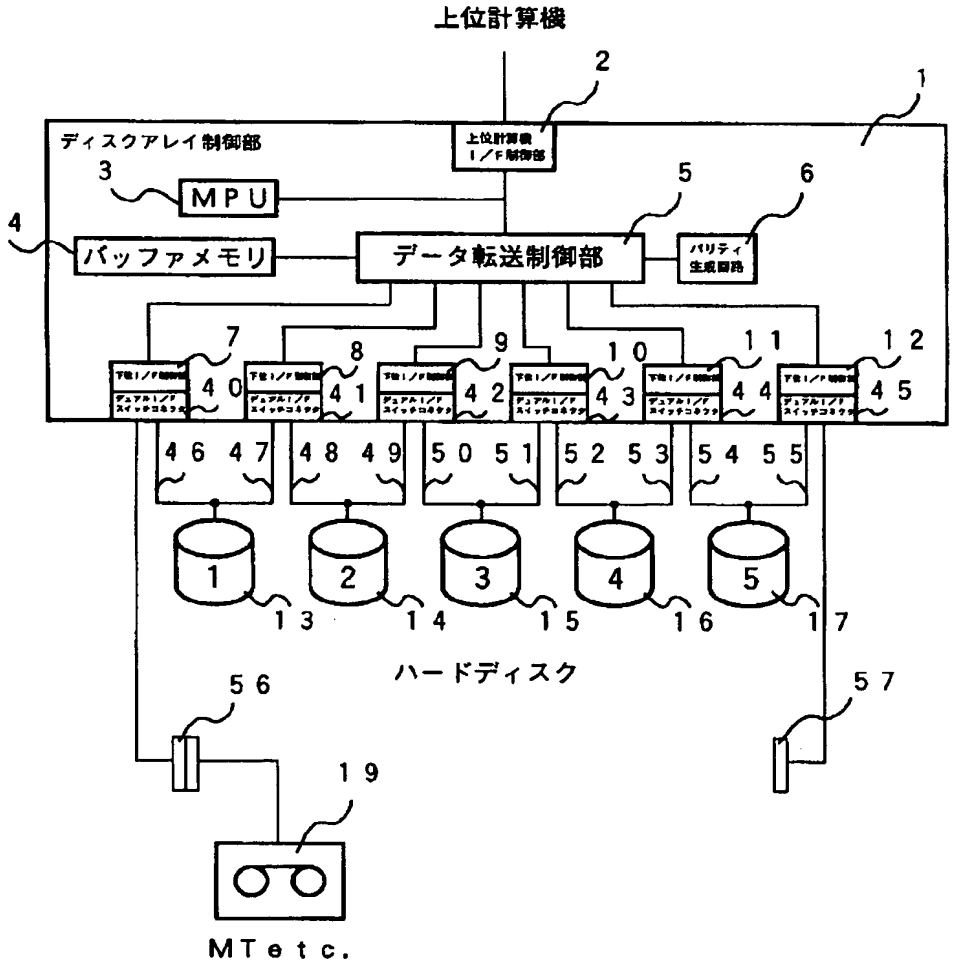
【図７】

図 7


200
ディスクアレイ装置をパワーオン
201
ディスクアレイ制御部がデュアルＩ／Ｆスイッチ
コネクタのＳＷＣＴＬ信号を偽にする
202
主経路を使用
203
上位計算機がコマンドを送出
204
リードか？
N
Y
205
ディスクアレイ制御部がデュアルＩ／Ｆスイッチ
コネクタのＷＲＥＮ信号を偽にする
206
ディスクアレイ制御部がデュアルＩ／Ｆスイッチ
コネクタのＷＲＥＮ信号を真にする
207
ハードディスクに対するリード／ライト
アクセス実行
208
下位Ｉ／Ｆ制御部、ケーブル
またはハードディスクに故障
発生か？
N
Y
209
ディスクアレイ制御部がデュアルＩ／Ｆスイッチ
コネクタのＳＷＣＴＬ信号を真にする
210
予備経路を使用
211
アクセス終了か？
N
Y
212
終了

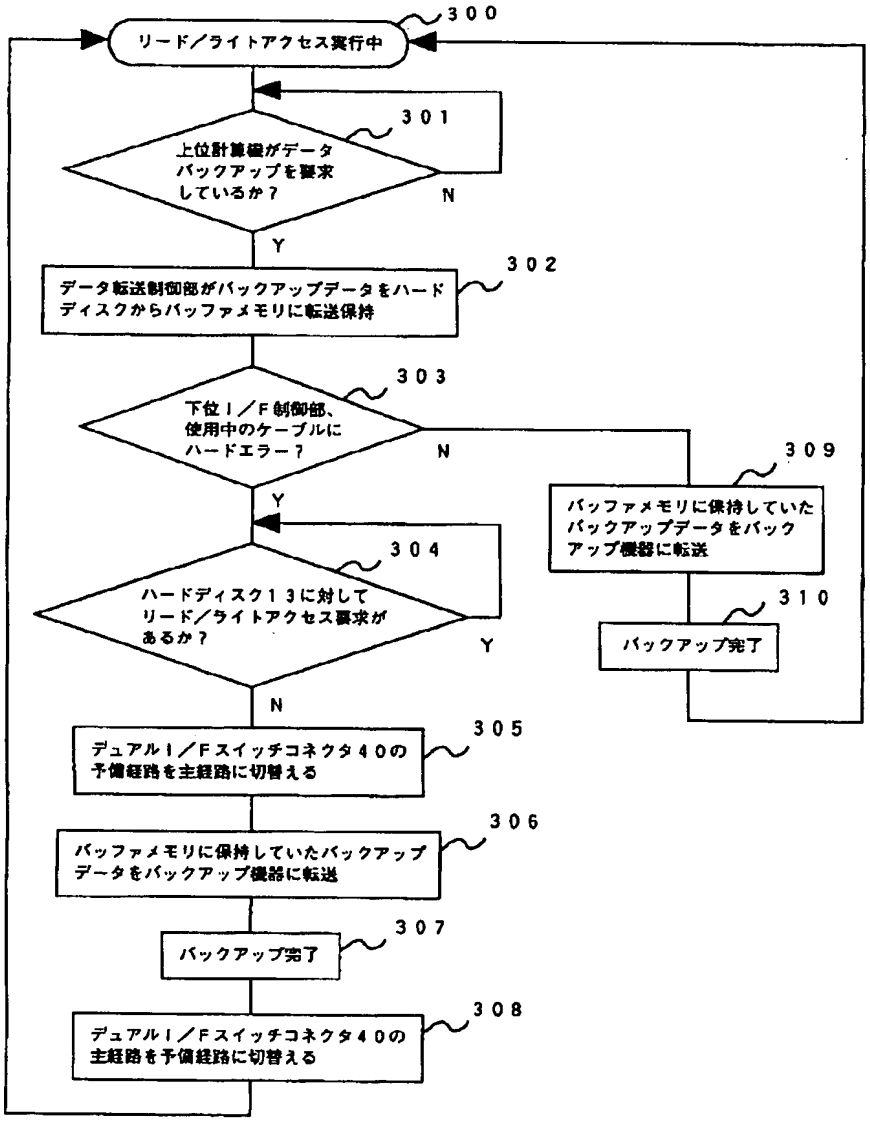
【図９】

図 ９


上位計算機
ディスクアレイ制御部
上位計算機 I／F制御部
MPU
バッファメモリ
データ転送制御部
パリティ生成回路
ハードディスク
MT etc.

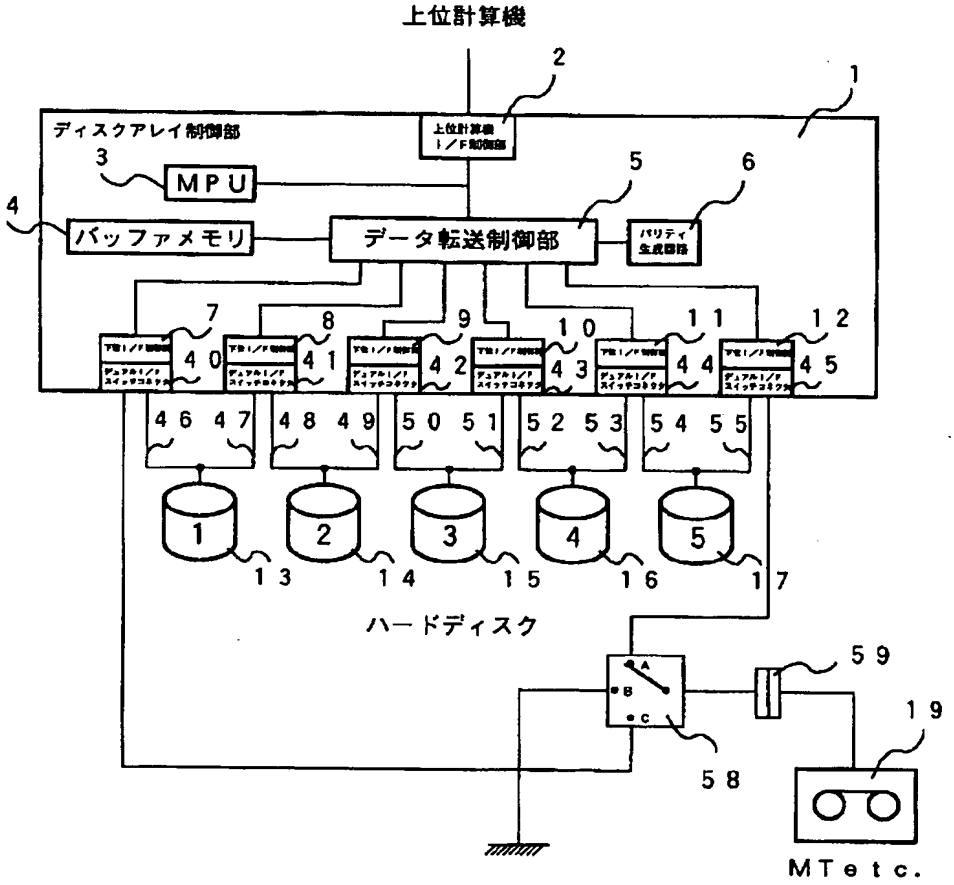
【図１０】

図 10


300
リード／ライトアクセス実行中
301
上位計算機がデータバックアップを要求しているか？
N
Y
302
データ転送制御部がバックアップデータをハードディスクからバッファメモリに転送保持
303
下位Ｉ／Ｆ制御部、使用中のケーブルにハードエラー？
N
Y
309
バッファメモリに保持していたバックアップデータをバックアップ機器に転送
310
バックアップ完了
304
ハードディスク１３に対してリード／ライトアクセス要求があるか？
Y
N
305
デュアルＩ／Ｆスイッチコネクタ４０の予備経路を主経路に切替える
306
バッファメモリに保持していたバックアップデータをバックアップ機器に転送
307
バックアップ完了
308
デュアルＩ／Ｆスイッチコネクタ４０の主経路を予備経路に切替える

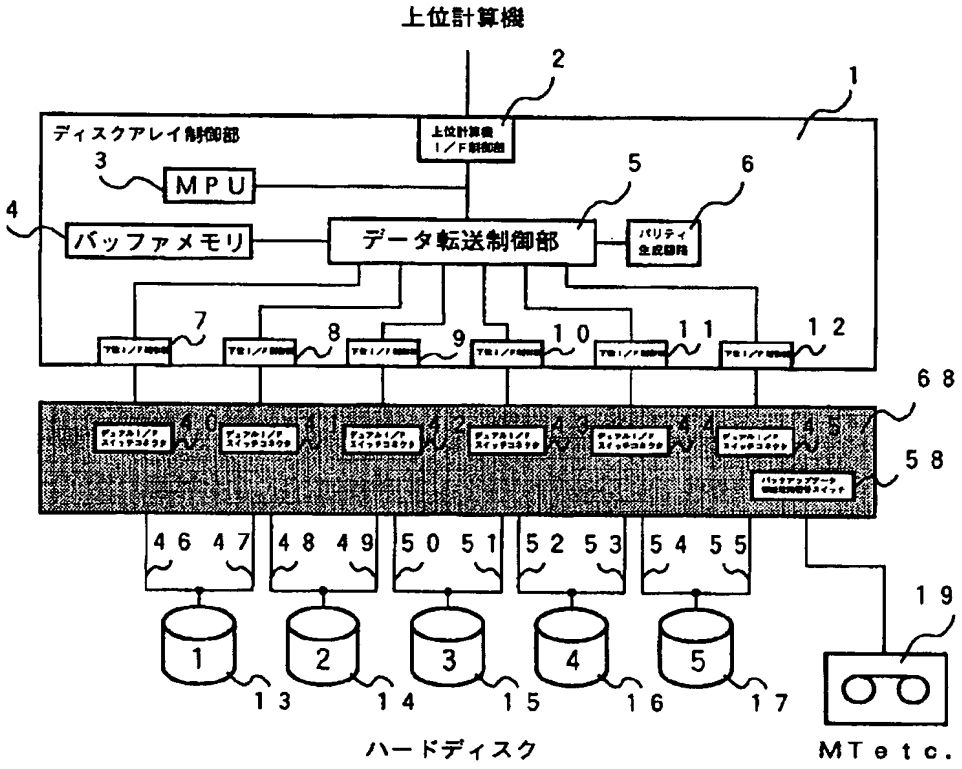
【図11】

図 11


上位計算機
2
1
ディスクアレイ制御部
3
MPU
4
バッファメモリ
データ転送制御部
5
6
パリティ生成回路
7
8
9
10
11
12
40
41
42
43
44
45
46 47
48 49
50 51
52 53
54 55
1
2
3
4
5
13
14
15
16
17
ハードディスク
59
19
58
A
B
C
MT etc.

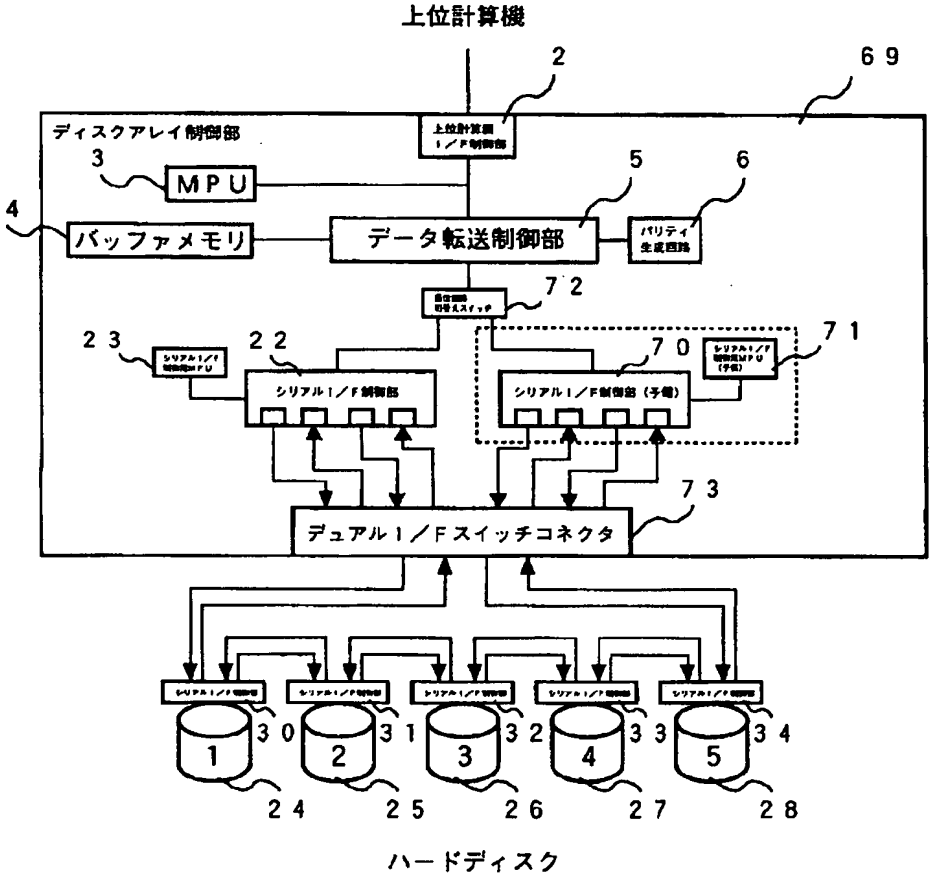
【図１２】

図 12


上位計算機
2
1
ディスクアレイ制御部
上位計算機
I／F制御部
3
MPU
5
6
4
バッファメモリ
データ転送制御部
パリティ
生成回路
7
8
9
10
11
12
68
58
46 47
48 49
50 51
52 53
54 55
19
1
2
3
4
5
13
14
15
16
17
ハードディスク
MT etc.

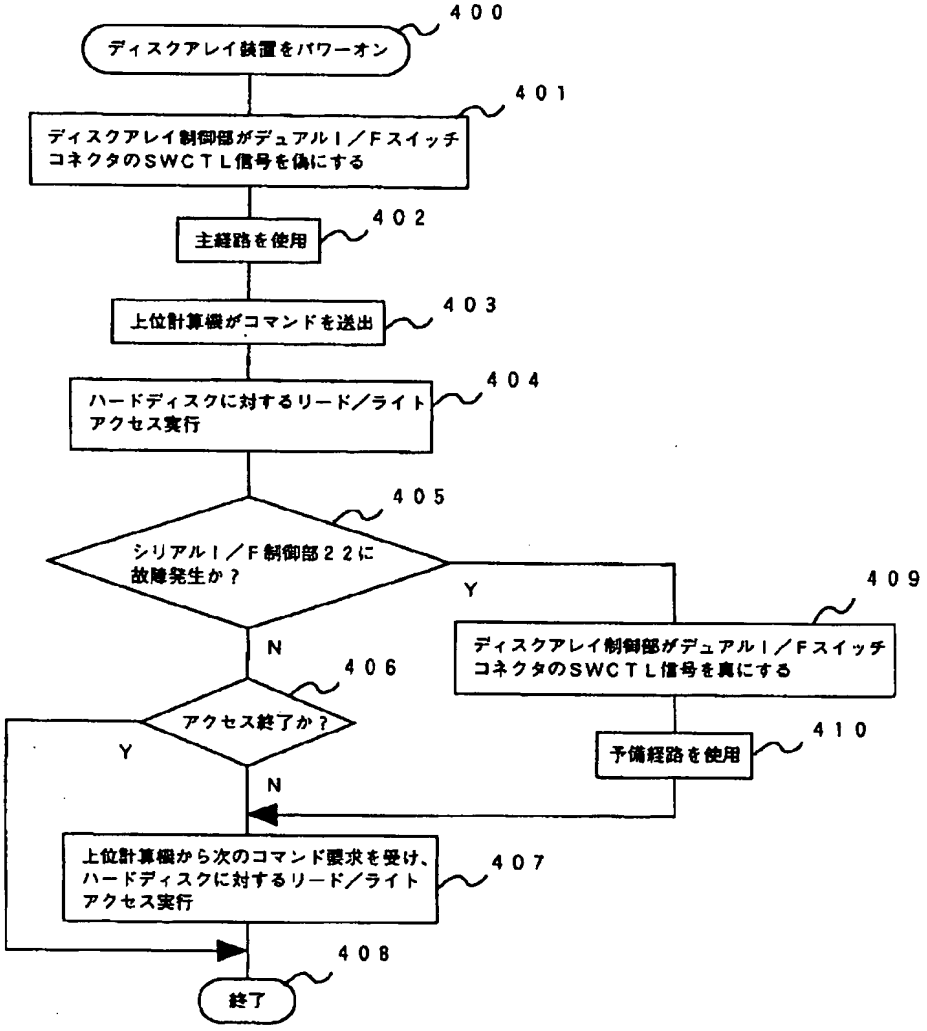
【図１３】

図 13


上位計算機
2
69
ディスクアレイ制御部
3
MPU
4
バッファメモリ
5
データ転送制御部
6
パリティ生成回路
72
23
22
70
71
シリアルI／F制御部
シリアルI／F制御部（予備）
73
デュアルI／Fスイッチコネクタ
1
2
3
4
5
30
31
32
33
34
24
25
26
27
28
ハードディスク

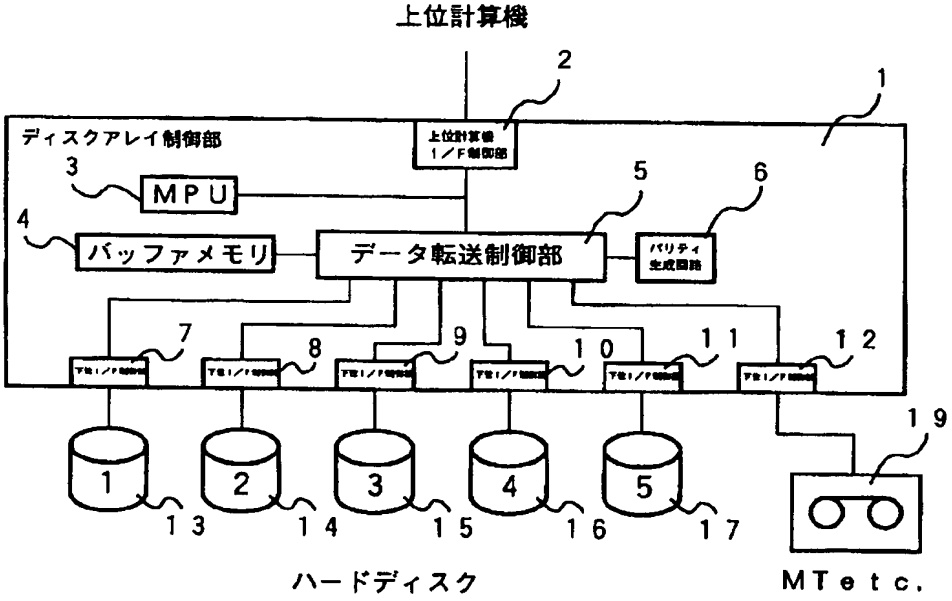
【図１５】

図 15


400
ディスクアレイ装置をパワーオン
401
ディスクアレイ制御部がデュアルＩ／ＦスイッチコネクタのＳＷＣＴＬ信号を偽にする
402
主経路を使用
403
上位計算機がコマンドを送出
404
ハードディスクに対するリード／ライトアクセス実行
405
シリアルＩ／Ｆ制御部２２に故障発生か？
Y
409
ディスクアレイ制御部がデュアルＩ／ＦスイッチコネクタのＳＷＣＴＬ信号を真にする
410
予備経路を使用
N
406
アクセス終了か？
Y
N
407
上位計算機から次のコマンド要求を受け、ハードディスクに対するリード／ライトアクセス実行
408
終了

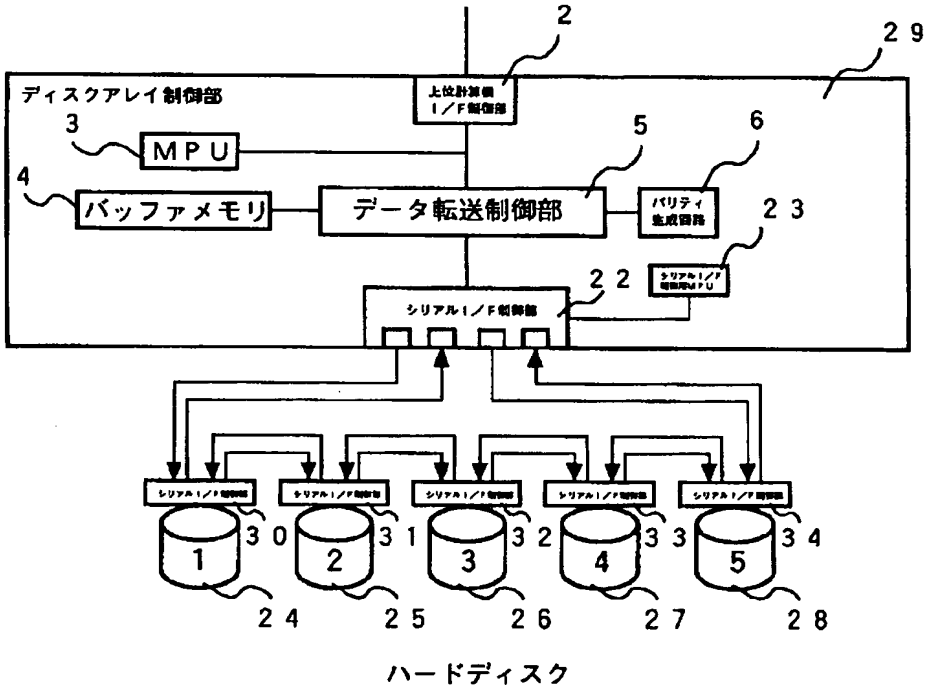
【図１６】

図 16


上位計算機
2
1
ディスクアレイ制御部
上位計算機I／F制御部
3
MPU
5
6
4
バッファメモリ
データ転送制御部
パリティ生成回路
7
8
9
10
11
12
19
1
2
3
4
5
13
14
15
16
17
ハードディスク
MT etc.

【図１８】

図 18



フロントページの続き

(72)発明者 本田 聖志
神奈川県川崎市麻生区王禅寺1099番地株式会社日立製作所システム開発研究所内

(72)発明者 岩崎 秀彦
神奈川県小田原市国府津2880番地株式会社日立製作所ストレージシステム事業部内